Econometric Theory, **18**, 2002, 744–775. Printed in the United States of America.
DOI: 10.1017.S0266466602183095

STRUCTURAL CHANGES AND SEEMINGLY UNIDENTIFIED STRUCTURAL EQUATIONS

IN CHOI
Hong Kong University of Science and Technology
and
University of Southern California

This paper introduces structural equations that do not satisfy the rank and/or order condition(s) for identification but still are identifiable. These equations are called seemingly unidentified structural equations. The key to the identifiability of these equations is that the right-hand-side endogenous variables undergo structural changes with respect to the exogenous and/or predetermined variables. To estimate the seemingly unidentified structural equations, this paper uses the classical minimum distance (MD) estimator and the principal components instrumental variables (PCIV) estimator. The PCIV estimator is different from conventional IV estimators for structural equations in that nonlinear functions of exogenous and/or predetermined variables are used as instruments. Simulation results comparing the estimator efficiency of the MD and PCIV estimators are reported in this paper. The results indicate that the MD and PCIV estimators are complementary to each other. The estimation methods proposed in this paper are applied to the Japanese export and GDP data to study the effect of export growth rate on that of GDP.

1. INTRODUCTION

Structural simultaneous equations have been one of the major research themes in econometrics since the 1950's,[1] although research on simultaneous equations has attracted relatively less attention these days partly as a result of a perceived failure in large macroeconometric models and partly because of criticisms of simultaneous equations modeling, following work by Lucas (1976) and Sims (1980). However, large-scale simultaneous equations models are cur-

I thank Don Andrews and two referees for very thoughtful comments and suggestions that I believe greatly improved this paper. Part of this paper was written while I was visiting the Department of Economics, University of Southern California. I am grateful to the faculty and staff of the department, especially Quang Vuong, Isabelle Perrigne, Cheng Hsiao, and Young Miller, for their support and hospitality. I also benefited from discussion with participants of econometrics seminars at ASU, UCLA, and UC Riverside. I thank Kangmo Choi for providing research assistance. This paper was financially supported by a research grant from Kookmin University. Some mathematical expressions in this paper were obtained with the help of the software *Scientific WorkPlace*™. Address correspondence to: In Choi, Department of Economics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong; e-mail: inchoi@ust.hk.

 © 2002 Cambridge University Press 0266-4666/02 $9.50

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

rently in widespread empirical use for forecasting and policy analyses throughout the world, especially in such nonacademic institutions as private banks and government economics departments, and small-scale simultaneous equations models are still used for academic research. Furthermore, the instrumental variables (IV) estimation methods used for structural equations are quite popular in both empirical and theoretical research. A recent surge of interest in IV estimation with weak instruments (e.g., Bekker, 1994; Bound, Jaeger, and Baker, 1995; Angrist and Krueger, 1995) demonstrates the relevance of the IV estimation methods to empirical analyses.

Prior to estimating the simultaneous equations, it is important to examine the identifiability of the equations. Phillips (1989), Choi and Phillips (1992, 1997), and Staiger and Stock (1997) focus on a single structural equation that is either unidentified or nearly unidentified and show that the conventional estimators for the structural equation do not have standard distributions, even asymptotically. Phillips (1989) and Choi and Phillips (1992, 1997) study the case where the rank condition for identification fails at any sample size, and Staiger and Stock (1997) examine the case where the equation is totally unidentified as sample size goes to infinity. In addition, Wang and Zivot (1998) and Zivot, Startz, and Nelson (1998) study the confidence interval of coefficient estimates for nearly unidentified structural equations by using the methods of Dufour (1997).

Phillips (1989), Choi and Phillips (1992, 1997), and Staiger and Stock (1997) prove how seriously unidentifiability affects the structural equation estimation and, thereby, demonstrate how difficult it is to build a proper structural equation. By contrast, this paper introduces structural equations that do not satisfy the rank and/or order condition(s) for identification but still are identifiable. These equations will be called seemingly unidentified structural equations, because the conventional identification conditions fail for these equations but the equations are still identifiable. The key to the identifiability of these equations is that the right-hand-side endogenous variables undergo structural changes with respect to the exogenous and/or predetermined variables. In the light of prevalent structural changes among economic variables, the structural equations studied here may be of much empirical relevance. Furthermore, the structural equations studied here can be used either for cross-section data or for time series.

Structural changes in a much simpler form than in this paper are also used for the IV estimation in Bekker and van der Ploeg (1999), though they do not make it explicit that their estimation methods utilize structural changes. But the structural changes in their work are simple mean shifts, which enables them to use dummy variables as instruments. Moreover, the model they study is in sharp contrast to ours in that it is identified in the conventional sense and has an infinite number of instruments as sample size grows to infinity. Angrist and Krueger (1995) is also related to this paper in that sample splitting is used for the IV estimation. But these authors' regression model is an identified one in the conventional sense, and the purpose of the sample splitting is to reduce bias, not to detect structural changes as in this paper.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

To estimate the seemingly unidentified structural equations, we use a minimum distance (MD) estimator. This estimator is obtained by using identifying relations for structural coefficients and reduced form parameter estimators.[2] The MD estimator is simple and intuitively appealing, and its efficiency can be considered within the framework of the MD estimation.

We also introduce a principal components instrumental variables (PCIV) estimator. The PCIV estimator uses principal components of the projected regressor matrix and is basically similar to the principal components OLS estimator except that instruments are used for the PCIV estimator. The PCIV estimator can estimate a linear combination of structural parameters under different data regimes. Therefore, each regime provides a linear combination of structural parameters, and the PCIV estimator for the structural parameters is defined as a solution of these equations.[3] A major characteristic of the PCIV estimator is that its asymptotic efficiency is not affected by the instruments chosen as long as the instruments are uncorrelated with the error terms of the structural equations.

The PCIV estimator is different from conventional IV estimators for structural equations in that nonlinear functions of exogenous and/or predetermined variables are used as instruments. This practice of using higher order moments of observed data has been prevalent in the literature of regressions with measurement error (cf. Lewbel, 1997; Dagenais and Dagenais, 1995). Using nonlinear functions of instruments in a nonparametric manner is also considered in the literature of nonlinear IV estimation (cf. Newey, 1990).

The MD and PCIV estimators are shown to have the same asymptotic distribution in a special case. But the asymptotic equivalence of the two estimators in general could not be derived in this paper.

Simulation results comparing the efficiency of the MD and PCIV estimators under known structural change points are reported. In addition, how misspecification of structural change points affects the estimators is studied by simulation. Simulation results for the MD and PCIV estimators using estimated structural change points are also presented.

The estimation methods proposed in this paper are applied to the Japanese export and gross domestic product (GDP) data to study the effect of export growth rate on that of GDP. The results reveal that exports had an accelerating multiplier effect on the GDP growth in Japan during the sample period 1960:Q2–1998:Q3.

The plan of the paper is as follows. Section 2 introduces seemingly unidentified structural equations and assumptions required for subsequent analyses. Section 3 studies the MD estimator and its asymptotic distributions. Section 4 introduces the PCIV estimator and derives its asymptotic distributions. Section 5 contains simulation results for estimator efficiency. Section 6 presents an empirical example. Section 7 contains a summary of the paper and some further remarks. All proofs are relegated to the Appendix.

The following notation will be used throughout the paper. For matrix A, $P_A = A(A'A)^{-1}A'$ and $Q_A = I - P_A$.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

2. THE MODELS AND ASSUMPTIONS

2.1. A Structural Equation under Rank Condition Failure

Suppose that a true structural relation is represented by the regression equation

$$y_{1i} = y_{2i}'\beta + z_{1i}'\gamma + u_i \qquad (i = 1, \ldots, N), \tag{1}$$

where y_{1i} and y_{2i} $(n \times 1)$ denote $n + 1$ endogenous variables, z_{1i} $(k_1 \times 1)$ is a vector of k_1 exogenous and/or predetermined variables, and u_i is a disturbance term.

Letting z_{2i} $(k_2 \times 1)$ be a vector of exogenous and/or predetermined variables excluded from equation (1), write the reduced form equations for the structural equation (1) as

$$y_{1i} = \pi_1' z_{1i} + \pi_2' z_{2i} + v_{1i} \tag{2}$$

and

$$y_{2i} = \Pi_1' z_{1i} + \Pi_2' z_{2i} + v_{2i}. \tag{3}$$

Letting $e_i' = [u_i \quad v_{2i}']$, assume

Assumption 1. $E(e_i e_k') = I(i = k)\Sigma$, where $I(\cdot)$ is an indicator function, $\Sigma = \begin{bmatrix} \sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{bmatrix}$, and $\Sigma_{uv} \neq 0$.

Under this assumption, e_i are uncorrelated, and y_{2i} and u_i are correlated for all i.

The identifying relations connecting the parameters of the structural equation (1) and the reduced form equations (2) and (3) are

$$\pi_1 - \Pi_1\beta = \gamma,$$

$$\pi_2 - \Pi_2\beta = 0.$$

Thus, when the rank condition

$$\operatorname{rank}(\Pi_2) = n \qquad (k_2 \geq n), \tag{4}$$

holds, the structural parameter vector $(\beta', \gamma')'$ is identifiable in the sense that the structural parameters can be recovered from the reduced form parameters, which are almost always estimable.

But it is assumed here that the following polar opposite case of the rank condition (4) holds.[4]

Assumption 2. $\Pi_2 = 0$.

Under this assumption and specification (3), the coefficient vector β is totally unidentified. Moreover, the coefficient vector γ is also unidentified under this assumption unless $\Pi_1 = 0$. Under Assumption 2, Phillips (1989) and Choi and Phillips (1992) show that the IV estimators using z_{1i} and z_{2i} as instruments are

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

inconsistent. The main reason for this is that z_{2i} is not a valid instrument for y_{2i} under Assumption 2.

However, in addition to Assumption 2, we make the following assumption.

Assumption 3.

$\Pi_1 = \Pi_{11}$ when $i \in \mathcal{G}_1$ and $\Pi_1 = \Pi_{12}$ when $i \in \mathcal{G}_2$,

where $\Pi_{11} \neq \Pi_{12}$, $\mathcal{G}_1 \cup \mathcal{G}_2 = \{1, \ldots, N\}$, $\mathcal{G}_1 \cap \mathcal{G}_2 = \varnothing$, and the numbers of elements of $\mathcal{G}_1$ and $\mathcal{G}_2$ are N_1 and N_2, respectively.

Under Assumptions 2 and 3,

$$y_{2i} = \Pi_{11}' z_{1i} + v_{2i}, \qquad i \in \mathcal{G}_1, \tag{5}$$

$$y_{2i} = \Pi_{12}' z_{1i} + v_{2i}, \qquad i \in \mathcal{G}_2.$$

That is, the endogenous variables y_{2i} undergo a structural change with respect to their relation to the exogenous and/or predetermined variables z_{1i} under Assumptions 2 and 3.

The identifying relations connecting the structural equation (1) and the reduced form equations (2) and (5) under Assumption 3 are

$$\pi_{11} - \Pi_{11}\beta = \gamma, \tag{6}$$

$$\pi_{12} - \Pi_{12}\beta = \gamma, \tag{7}$$

where π_{1j} denotes the coefficient π_1 under $i \in \mathcal{G}_j$. These equations show that the structural parameters are identifiable even under the rank condition failure, because equations (6) and (7) can be solved with respect to the parameter vectors β and γ when the following assumption is satisfied.

Assumption 4. $\operatorname{rank}\begin{pmatrix} I_{k_1} & \Pi_{11} \\ I_{k_1} & \Pi_{12} \end{pmatrix} = n + k_1$.

Note that Assumption 4 implicitly requires $n \leq k_1$.

The key element in identifying the structural equation (1) is that the reduced form equation for the endogenous variable y_{2i} undergoes a structural change, whereas the structural equation does not. Structural equations are usually based on economic theory or reasoning that is assumed to hold true for any economic agent and in any time period. Economic theory seldom predicts structural changes. Thus, assuming parameter constancy for the structural equation may not be too unreasonable unless significant and widespread social changes (e.g., changes in economic system, revolutions, etc.) are known to occur. However, reduced form equations merely signify data generating processes for which structural changes are often observed. A well known example is Perron (1989). Notwithstanding these arguments, testing whether the identifiability conditions—absence of structural changes for the structural equation and the presence of a structural change in the reduced form equation—hold for a given data set is important in empirical practices and, therefore, needs to be pursued.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

We have considered the structural equation under the rank condition failure, but this is also related to a general IV estimation problem. Write equations (1) and (3) under Assumption 2 as

$$y_{1i} = w_i'\delta + u_i,$$

where $w_i = [y_{2i}', z_{1i}']', \delta = [\beta', \gamma']'$, and

$$\begin{pmatrix} y_{2i} \\ z_{1i} \end{pmatrix} = \begin{pmatrix} \Pi_1' \\ I_{k_1} \end{pmatrix} z_{1i} + \begin{pmatrix} v_{2i} \\ 0 \end{pmatrix}$$

or equivalently

$$w_i = \Xi' z_{1i} + v_i.$$

Written in this way, equations (1) and (3) under Assumption 2 denote the IV regression with insufficient number of instruments. It is inferred from the discussions up to now that the problem of insufficient number of instruments in the IV estimation can be overcome where there is a structural change in the relation between the endogenous variables and instruments. For this reason, estimation theory that will be developed in subsequent sections can equally be applied to IV estimation with insufficient number of instruments.

A time series example of a seemingly unidentified structural equation under the rank condition failure is

$$y_{1i} = \beta y_{2i} + \gamma y_{2(i-1)} + u_i \tag{8}$$

with

$$\begin{aligned} y_{2i} &= \Pi_{11} y_{2(i-1)} + v_{2i}, \qquad |\Pi_{11}| < 1, \qquad i \in \mathcal{G}_1, \\ y_{2i} &= \Pi_{12} y_{2(i-1)} + v_{2i}, \qquad |\Pi_{12}| < 1, \qquad i \in \mathcal{G}_2, \end{aligned} \tag{9}$$

and

$$\operatorname{rank} \begin{pmatrix} 1 & \Pi_{11} \\ 1 & \Pi_{12} \end{pmatrix} = 2.$$

Equations (8) and (9) signify a typical distributed lag model where the right-hand-side variable of the structural equation follows an AR(1) process. One may usually be interested in the value of $\beta + \gamma$, which is called the long-run multiplier. When there is no structural change in the AR(1) process, the structural equation (8) is unidentifiable. But a structural change as in equation (9) makes the structural equation identifiable.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

2.2. A Structural Equation under Order Condition Failure

Consider a structural equation

$$y_{1i} = y_{2i}'\beta + z_{1i}'\gamma + z_{2i}'\rho + u_i, \tag{10}$$

where y_{2i}, z_{1i}, and u_i are as defined in Section 2.1 and z_{2i} $(k_2 \times 1)$ is a vector of exogenous and/or predetermined variables. We treat this equation separately from equation (1) because the regressor z_{2i} is included in the regression equation (10) regardless of the true value of the parameter vector ρ.

The identifying relations connecting the parameters of the structural equation (10) and the reduced form equations (2) and (3) are

$$\pi_1 - \Pi_1\beta = \gamma,$$

$$\pi_2 - \Pi_2\beta = \rho. \tag{11}$$

The condition $\rho = 0$ is a necessary condition for the identification of equation (10) and has been called the order condition for identification.[5] When this condition is violated, it is obvious that the IV estimator using z_{1i} and z_{2i} as instruments is inconsistent. However, it is assumed in this paper that there are no restrictions on ρ as in the following assumption.

Assumption 5. $\rho \in R^{k_2}$.

To make the structural parameters identified, however, we assume that Assumption 3 and the following assumption hold.

Assumption 6. $\operatorname{rank}\begin{pmatrix} I_{k_1} & 0 & \Pi_{11} \\ I_{k_1} & 0 & \Pi_{12} \\ 0 & I_{k_2} & \Pi_2 \end{pmatrix} = n + k_1 + k_2.$

Under this assumption and Assumption 3, equations (6), (7), and (11) can be solved with respect to the structural parameters β, γ, and ρ. For simplicity, it is assumed here that the reduced form parameter Π_2 does not undergo structural changes. Extending the results of this paper to the case where Π_2 changes over regimes $\mathcal{G}_1$ and $\mathcal{G}_2$ is straightforward. In addition, Assumption 6 implicitly requires $n \leq k_1$.

Assumption 6 replaces Assumption 4, so that it is not required here anymore. Furthermore, Π_2 may take any value as long as it satisfies Assumption 6, and hence Assumption 2 is not needed in this section. In summary, this section requires Assumptions 1, 3, 5, and 6.

Equation (10) with the reduced form equation (3) may be interpreted as an IV regression model with insufficient number of instruments, as in Section 2.1. Thus, the IV regression model can be identified when the endogenous variables in the model undergo a structural change with respect to a set of instru-

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

ments, the number of which is greater than or equal to the number of the endogenous variables.

As a time series example of a seemingly unidentified structural equation under the order condition failure, consider

$$y_{1i} = \beta y_{2i} + \gamma y_{2(i-1)} + \rho x_{1i} + u_i, \tag{12}$$

where x_{1i} is a strictly exogenous variable,

$$y_{2i} = \Pi_{11} y_{2(i-1)} + \Pi_2 x_{1i} + v_{2i}, \qquad |\Pi_{11}| < 1, \quad i \in \mathcal{G}_1, \tag{13}$$

$$y_{2i} = \Pi_{12} y_{2(i-1)} + \Pi_2 x_{1i} + v_{2i}, \qquad |\Pi_{12}| < 1, \quad i \in \mathcal{G}_2$$

and

$$\operatorname{rank}\begin{pmatrix} 1 & 0 & \Pi_{11} \\ 1 & 0 & \Pi_{12} \\ 0 & 1 & \Pi_2 \end{pmatrix} = 3.$$

Equation (12) is a typical structural equation that violates the order condition. However, with a structural change as in equation (13), the structural equation (12) would be identified. Parameter ρ may take any value in this example.

3. MINIMUM DISTANCE ESTIMATION

3.1. Minimum Distance Estimator

This section considers MD estimators for the structural parameters in equations (1) and (10). The objective function for the MD estimation of the parameter vector $\delta = (\beta', \gamma')'$ is a weighted product of the sample version of the identifying relations (6) and (7). That is, the MD estimator of the coefficient vector $\delta = (\beta', \gamma')'$ in the regression equation (1) is obtained by minimizing the objective function

$$R(\delta) = (\hat{\pi} - \hat{\Pi}\delta)' \Upsilon_{\delta}^{-1} (\hat{\pi} - \hat{\Pi}\delta)$$

with respect to the parameter vector δ, where Υ_{δ}^{-1} is a weight matrix. Here, letting $Z_{1j} = [z'_{1i}]_{i \in \mathcal{G}_j}, Y_{1j} = [y_{1i}]_{i \in \mathcal{G}_j}$, and $Y_{2j} = [y'_{2i}]_{i \in \mathcal{G}_j}$ $(j = 1,2)$,

$$\hat{\pi} = \begin{bmatrix} \hat{\pi}_{11} \\ \hat{\pi}_{12} \end{bmatrix}, \qquad \hat{\Pi} = \begin{bmatrix} \hat{\Pi}_{11} & I_{k_1} \\ \hat{\Pi}_{12} & I_{k_1} \end{bmatrix},$$

$$\hat{\pi}_{1j} = (Z'_{1j} Z_{1j})^{-1} Z'_{1j} Y_{1j}, \hat{\Pi}_{1j} = (Z'_{1j} Z_{1j})^{-1} Z'_{1j} Y_{2j}.$$

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

The best efficiency of the MD estimator is attained when[6]

$$\Upsilon_\delta = \frac{1}{\sigma_{uu}} E(\hat{\pi} - \hat{\Pi}\delta)(\hat{\pi} - \hat{\Pi}\delta)' = \begin{bmatrix} (Z_{11}'Z_{11})^{-1} & 0 \\ 0 & (Z_{12}'Z_{12})^{-1} \end{bmatrix}. \tag{14}$$

Note that the second equality in equation (14) follows from the relation $\hat{\pi} - \hat{\Pi}\delta = AU$, where

$$A = \begin{bmatrix} (Z_{11}'Z_{11})^{-1}Z_{11}' & 0 \\ 0 & (Z_{12}'Z_{12})^{-1}Z_{12}' \end{bmatrix} \tag{15}$$

and $U = [u_1, \ldots, u_N]'$. The resulting MD estimator of δ is written as

$$\hat{\delta}_{\mathrm{MD}} = (\hat{\Pi}'\Upsilon_{\delta}^{-1}\hat{\Pi})^{-1}\hat{\Pi}'\Upsilon_{\delta}^{-1}\hat{\pi}. \tag{16}$$

Analogously, letting $Z_{2j} = [z_{2i}']_{i\in\mathcal{G}_j}$, $Z_1 = [z_{1i}']_{i=1,\ldots,N}$, $Z_2 = [z_{2i}']_{i=1,\ldots,N}$, $Y_1 = [y_{1i}]_{i=1,\ldots,N}$, and $Y_2 = [y_{2i}']_{i=1,\ldots,N}$, the MD estimator of the parameter vector $\phi = (\beta', \gamma', \rho')'$ for the regression equation (10) is defined as

$$\bar{\phi}_{\mathrm{MD}} = (\bar{\Pi}'\Upsilon_{\phi}^{-1}\bar{\Pi})^{-1}\bar{\Pi}'\Upsilon_{\phi}^{-1}\bar{\pi}, \tag{17}$$

where

$$\bar{\pi} = \begin{bmatrix} \bar{\pi}_{11} \\ \bar{\pi}_{12} \\ \bar{\pi}_2 \end{bmatrix}, \qquad \bar{\Pi} = \begin{bmatrix} \bar{\Pi}_{11} & I_{k_1} & 0 \\ \bar{\Pi}_{12} & I_{k_1} & 0 \\ \bar{\Pi}_2 & 0 & I_{k_2} \end{bmatrix}, \qquad \Upsilon_\phi = \begin{bmatrix} \Upsilon_{\phi 11} & \Upsilon_{\phi 12} & \Upsilon_{\phi 13} \\ \Upsilon_{\phi 12}' & \Upsilon_{\phi 22} & \Upsilon_{\phi 23} \\ \Upsilon_{\phi 13}' & \Upsilon_{\phi 23}' & \Upsilon_{\phi 33} \end{bmatrix},$$

$$\bar{\pi}_{1j} = (Z_{1j}'Q_{Z_{2j}}Z_{1j})^{-1}Z_{1j}'Q_{Z_{2j}}Y_{1j} \qquad (j = 1,2),$$

$$\bar{\pi}_2 = (Z_2'Q_{Z_1}Z_2)^{-1}Z_2'Q_{Z_1}Y_1,$$

$$\bar{\Pi}_{1j} = (Z_{1j}'Q_{Z_{2j}}Z_{1j})^{-1}Z_{1j}'Q_{Z_{2j}}Y_{2j} \qquad (j = 1,2),$$

$$\bar{\Pi}_2 = (Z_2'Q_{Z_1}Z_2)^{-1}Z_2'Q_{Z_1}Y_2,$$

$$\Upsilon_{\phi 11} = (Z_{11}'Q_{Z_{21}}Z_{11})^{-1}, \qquad \Upsilon_{\phi 22} = (Z_{12}'Q_{Z_{22}}Z_{12})^{-1}, \qquad \Upsilon_{\phi 33} = (Z_2'Q_{Z_1}Z_2)^{-1},$$

$$\Upsilon_{\phi 12} = 0, \qquad \Upsilon_{\phi 13} = \Upsilon_{\phi 23} = -(Z_1'Z_1)^{-1}Z_1'Z_2(Z_2'Q_{Z_1}Z_2)^{-1}.$$

It follows from the relation $\bar{\pi} = \bar{\Pi}\phi + BU$, where

$$B = \begin{bmatrix} (Z_{11}'Q_{Z_{21}}Z_{11})^{-1}Z_{11}'Q_{Z_{21}}[I_{N_1}, 0] \\ (Z_{12}'Q_{Z_{22}}Z_{12})^{-1}Z_{12}'Q_{Z_{22}}[0, I_{N_2}] \\ (Z_2'Q_{Z_1}Z_2)^{-1}Z_2'Q_{Z_1} \end{bmatrix}, \tag{18}$$

that $\Upsilon_\phi = BB'$.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

3.2. Asymptotic Distribution Theory

This section reports asymptotic distributions of the MD estimators introduced in the last section. For the asymptotic distributions of the MD estimators defined in (16) and (17), the following assumption is required.

Assumption 7. As $N, N_j \to \infty$ $(j = 1,2)$,

(a) $(1/N_j)Z_{1j}'Z_{1j} \xrightarrow{p} M_{z_{1j}z_{1j}}$, where $M_{z_{1j}z_{1j}}$ is a positive definite nonrandom matrix;
(b)

$$\frac{1}{N_j} Z^{j\prime} Z^j \xrightarrow{p} \begin{bmatrix} M_{z_{1j}z_{1j}} & M_{z_{1j}z_{2j}} \\ M_{z_{2j}z_{1j}} & M_{z_{2j}z_{2j}} \end{bmatrix} \begin{matrix} k_1 \\ k_2 \end{matrix},$$

where $Z^j = [Z_{1j}, Z_{2j}]$ and the probability limit of $(1/N_j)Z^{j\prime}Z^j$ is a positive definite nonrandom matrix and partitioned conformably;
(c) $(1/N_j)Z^{j\prime}V_{1j} \xrightarrow{p} 0$ and $(1/N_j)Z^{j\prime}V_{2j} \xrightarrow{p} 0$, where $V_{1j} = [v_{1i}]_{i\in\mathcal{G}_j}$ and $V_{2j} = [v_{2i}']_{i\in\mathcal{G}_j}$;
(d) $(N_j/N) \to d_j$ where $0 < d_j < 1$;
(e) For $K = A$ or B,

$$\sqrt{N}KU \xrightarrow{d} N(0, \sigma_{uu} \operatorname{plim}_{N\to\infty} NKK'),$$

where the matrices A and B are given in equations (15) and (18), respectively.

Assumption 7 follows from the law of large numbers and the central limit theorem when specific conditions are given to the involved random variables.

The asymptotic distributions of the minimum distance estimators $\hat{\delta}_{\mathrm{MD}}$ and $\bar{\phi}_{\mathrm{MD}}$ are reported in the following theorem.

THEOREM 1. *(a) Under Assumptions 1–4 and 7, as* $N, N_j \to \infty$ $(j = 1,2)$,

$$\sqrt{N}(\hat{\delta}_{\mathrm{MD}} - \delta) \xrightarrow{d} N(0, \sigma_{uu}\Omega^{-1}),$$

where

$$\Omega = \begin{bmatrix} \Pi_{11} & I_{k_1} \\ \Pi_{12} & I_{k_1} \end{bmatrix}' \begin{bmatrix} \frac{1}{d_1} M_{z_{11}z_{11}}^{-1} & 0 \\ 0 & \frac{1}{d_2} M_{z_{12}z_{12}}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \Pi_{11} & I_{k_1} \\ \Pi_{12} & I_{k_1} \end{bmatrix}.$$

(b) Assume that the matrix D defined subsequently is nonsingular. Under Assumptions 1, 3, and 5–7, as $N, N_j \to \infty$ $(j = 1,2)$,

$$\sqrt{N}(\bar{\phi}_{\mathrm{MD}} - \delta) \xrightarrow{d} N(0, \sigma_{uu}\bar{\Omega}^{-1}),$$

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

where

$$\bar{\Omega} = \begin{bmatrix} \Pi_{11} & I_{k_1} & 0 \\ \Pi_{12} & I_{k_1} & 0 \\ \Pi_2 & 0 & I_{k_2} \end{bmatrix}' D^{-1} \begin{bmatrix} \Pi_{11} & I_{k_1} & 0 \\ \Pi_{12} & I_{k_1} & 0 \\ \Pi_2 & 0 & I_{k_2} \end{bmatrix},$$

$$D = \begin{bmatrix} \frac{1}{d_1} M^{-1}_{z_{11}\bullet z_{21}} & 0 & -M^{-1}_{z_1 z_1} M_{z_1 z_2} M^{-1}_{z_2 \bullet z_1} \\ 0 & \frac{1}{d_2} M^{-1}_{z_{12}\bullet z_{22}} & -M^{-1}_{z_1 z_1} M_{z_1 z_2} M^{-1}_{z_2 \bullet z_1} \\ -M^{-1}_{z_2 \bullet z_1} M_{z_2 z_1} M^{-1}_{z_1 z_1} & -M^{-1}_{z_2 \bullet z_1} M_{z_2 z_1} M^{-1}_{z_1 z_1} & M^{-1}_{z_2 \bullet z_1} \end{bmatrix},$$

$M_{a\bullet b} = M_{aa} - M_{ab} M^{-1}_{bb} M_{ba}$,

and $\begin{bmatrix} M_{z_1 z_1} & M_{z_1 z_2} \\ M_{z_2 z_1} & M_{z_2 z_2} \end{bmatrix}$ *is the probability limit of* $(1/N)Z'Z$ *with* $Z = [Z_1, Z_2]$.

Remarks

(a) This theorem shows that the MD estimators are consistent and normally distributed in the limit.
(b) When Assumptions 4 and 6 do not hold, the variance-covariance matrices in Theorem 1 become singular, which indicates the lack of identifiability.
(c) In the special case $n = k_1 = 1$, the variance-covariance matrix Ω becomes

$$\sigma_{uu} \begin{bmatrix} \dfrac{d_1 M_{z_{11} z_{11}} + d_2 M_{z_{12} z_{12}}}{d_1 M_{z_{11} z_{11}} d_2 M_{z_{12} z_{12}} (\Pi_{12} - \Pi_{11})^2} & -\dfrac{\Pi_{11} d_1 M_{z_{11} z_{11}} + \Pi_{12} d_2 M_{z_{12} z_{12}}}{d_1 M_{z_{11} z_{11}} d_2 M_{z_{12} z_{12}} (\Pi_{12} - \Pi_{11})^2} \\ -\dfrac{\Pi_{11} d_1 M_{z_{11} z_{11}} + \Pi_{12} d_2 M_{z_{12} z_{12}}}{d_1 M_{z_{11} z_{11}} d_2 M_{z_{12} z_{12}} (\Pi_{12} - \Pi_{11})^2} & \dfrac{\Pi^2_{11} d_1 M_{z_{11} z_{11}} + \Pi^2_{12} d_2 M_{z_{12} z_{12}}}{d_1 M_{z_{11} z_{11}} d_2 M_{z_{12} z_{12}} (\Pi_{12} - \Pi_{11})^2} \end{bmatrix}. \quad \textbf{(19)}$$

Several observations emerge from the asymptotic variance-covariance matrix (19). First, the asymptotic variances of $\hat{\beta}_{\mathrm{MD}}$ and $\hat{\gamma}_{\mathrm{MD}}$ decrease as $(\Pi_{12} - \Pi_{11})^2$ increases. That is, more abrupt structural changes yield higher asymptotic efficiency. Second, the variance and covariance formulae for $\hat{\beta}_{\mathrm{MD}}$ and $\hat{\gamma}_{\mathrm{MD}}$ break down when $\Pi_{11} = \Pi_{12}$, which indicates the lack of identifiability.

4. PRINCIPAL COMPONENTS IV ESTIMATION

4.1. Basic Ideas and Assumptions

The coefficients for the structural equation under regime j ($j = 1,2$),

$$y_{1i} = y'_{2i}\beta + z'_{1i}\gamma + u_i, \qquad i \in \mathcal{G}_j \quad \textbf{(20)}$$

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

with

$$y_{2i} = \Pi_{1j}' z_{1i} + v_{2i}, \qquad i \in \mathcal{G}_j,$$

cannot be estimated by the two-stage least squares estimation as discussed in Phillips (1989) and Choi and Phillips (1992). The two-stage least squares estimation fails mainly because $\hat{y}_{2i}$ $(= \hat{\Pi}_{1j}' z_{1i}$; $\hat{\Pi}_{1j}$ is the ordinary least squares [OLS] estimator) and z_{1i} are collinear. But when there exists multicollinearity among regressors, it is well known[7] that the principal component regression can still estimate a linear combination of the coefficient vector. This implies that we can estimate two linear combinations of the structural coefficient vector from the two data regimes $\mathcal{G}_1$ and $\mathcal{G}_2$ and use them to estimate the structural coefficient vector.

In using the principal component regression for the regression equation (20), there are two problems that deserve our attention. First, the regressor y_{2i} and the disturbance term u_i are correlated. Second, z_{2i} may not be available for use as instruments because of Assumption 2. To resolve these problems, we will use the principal component regression using instrument variables, which we call principal component instrumental variables (PCIV) regression. As instruments, nonlinear functions of z_{1i}, $H(z_{1i}) = h_i$,[8] are used because other instruments may not be available in practice.[9] It is assumed that the number of elements of the vector h_i is greater than or equal to n (the number of elements in vector y_{2i}) for the condition of sufficient number of instruments to be satisfied. In the literature of regressions with measurement error, higher order moments of observed data have long been used as instruments.[10] In this paper, this practice is adopted to estimate structural equations.

In the same manner, the PCIV regression method can be used to estimate the structural coefficient vector for equation (10). In this case, we use $F(z_{1i}, z_{2i}) = f_i$ as instruments for y_{2i}, where $F(\cdot,\cdot)$ is a nonlinear function and the number of elements of f_i is greater than or equal to that of y_{2i}.

Now we introduce assumptions required for the development of distribution theory. The instrument matrices that we will use for equations (1) and (10) are

$$G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}, \qquad G_j = [h_i', z_{1i}']_{i \in \mathcal{G}_j} \quad \text{and} \quad \bar{G} = \begin{bmatrix} \bar{G}_1 \\ \bar{G}_2 \end{bmatrix},$$

$$\bar{G}_j = [f_i', z_{1i}', z_{2i}']_{i \in \mathcal{G}_j} \qquad (j = 1,2),$$

respectively, for which the following assumption is made for the PCIV estimation.

Assumption 8. As $N, N_j \to \infty$ $(j = 1,2)$,

(a) $(1/N_j) G_j' G_j \xrightarrow{p} M_{G_j G_j} = \begin{bmatrix} M_{h_j h_j} & M_{h_j z_{1j}} \\ M_{z_{1j} h_j} & M_{z_{1j} z_{1j}} \end{bmatrix}$, where $M_{G_j G_j}$ is a positive definite nonrandom matrix and partitioned conformably;

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

(b) $(1/N_j)\bar{G}_j'\bar{G}_j \xrightarrow{p} M_{\bar{G}_j\bar{G}_j} = \begin{bmatrix} M_{f_jf_j} & M_{f_jz_{1j}} & M_{f_jz_{2j}} \\ M_{z_{1j}f_j} & M_{z_{1j}z_{1j}} & M_{z_{1j}z_{2j}} \\ M_{z_{2j}f_j} & M_{z_{2j}z_{1j}} & M_{z_{2j}z_{2j}} \end{bmatrix}$, where $M_{\bar{G}_j\bar{G}_j}$ is a positive definite nonrandom matrix and partitioned conformably;
(c) $(1/N_j)V_{2j}'L \xrightarrow{p} 0$, where $L = G_j$ or $\bar{G}_j$;
(d) $(N_j/N) \to d_j$, where $0 < d_j < \infty$;
(e) for $L = G_j$ or $\bar{G}_j$,

$$\frac{1}{\sqrt{N_j}} L'U \xrightarrow{d} N\left(0, \sigma_{uu}\text{plim}_{N_j\to\infty}\frac{1}{N_j}L'L\right).$$

Parts (a) and (b) of Assumption 8 essentially imply that the law of large numbers holds for the products of the instrument matrices and that their probability limits are positive definite. Note that the instrument matrices are allowed to be stochastic for our asymptotic analysis. Part (c) follows when more specific conditions are given on matrix V_{2j}. For example, v_{2i} may be either vector martingale difference sequence or vector independent and identically distributed (i.i.d.) sequence. Part (e) follows from the central limit theorem with more specific conditions on the errors and instruments.

4.2. Principal Components IV Estimator

This section formally introduces the PCIV estimator for the seemingly unidentified structural equations in Section 2. The principal components OLS regression requires knowing how many nonzero eigenvalues are present in the product of a regressor matrix. Likewise, the PCIV estimator applied to each regime of models (1) and (10) will require the number of nonzero eigenvalues for the product of a projected regressor matrix. The following lemma will be used to find such numbers.

LEMMA 1.

(a) Under Assumptions 2 and 8, $\text{rank}(M_{W_jG_j}M_{G_jG_j}^{-1}M_{G_jW_j}) = k_1$, *where* $M_{W_jG_j}M_{G_jG_j}^{-1}M_{G_jW_j} = \text{plim}_{N_j\to\infty}(1/N_j)W_j'P_{G_j}W_j$ *and* $W_j = [y_{2i}', z_{1i}']_{i\in\mathcal{G}_j}$ $(j = 1,2)$.
(b) Under Assumption 8, $\text{rank}(M_{\overline{W}_j\bar{G}_j}M_{\bar{G}_j\bar{G}_j}^{-1}M_{\bar{G}_j\overline{W}_j}) = k_1 + k_2$, *where* $M_{\overline{W}_j\bar{G}_j}M_{\bar{G}_j\bar{G}_j}^{-1}M_{\bar{G}_j\overline{W}_j} = \text{plim}_{N_j\to\infty}(1/N_j)\overline{W}_j'P_{\bar{G}_j}\overline{W}_j$ *and* $\overline{W}_j = [y_{2i}', z_{1i}', z_{2i}']_{i\in\mathcal{G}_j}$ $(j = 1,2)$.

To introduce the PCIV estimator, write equation (20) in matrix notation as

$$\begin{aligned} Y_{1j} &= W_j\delta + U_j, \qquad (U_j = [u_i]_{i\in\mathcal{G}_j}) \\ &= W_jS_jS_j'\delta + U_j \\ &= W_jS_{j1}\theta_{j1} + W_jS_{j2}\theta_{j2} + U_j \\ &= X_{1j}\theta_{j1} + X_{2j}\theta_{j2} + U_j, \end{aligned} \tag{21}$$

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

where $S_j = \begin{bmatrix} k_1 & n \\ S_{j1} & S_{j2} \end{bmatrix}$ is a matrix whose columns are orthonormal eigenvectors of $W_j' P_{G_j} W_j$, $\theta_{jm} = S_{jm}'\delta$ $(m = 1,2)$ and $X_{mj} = W_j S_{jm}$. Here, the eigenvectors are ordered such that

$$S_j' W_j' P_{G_j} W_j S_j = \text{diag}[\lambda_{j1}, \ldots, \lambda_{j(n+k_1)}] \quad \text{and} \quad \lambda_{j1} \geq \cdots \geq \lambda_{j(n+k_1)}.$$

Thus, vectors in the matrix S_{j1} correspond to the k largest eigenvalues $\lambda_{j1}, \ldots, \lambda_{jk_1}$.

Using part (a) of Lemma 1, the PCIV estimators of θ_{j1} and θ_{j2} using G_j as an instrument matrix are defined as

$$\hat{\theta}_{j1\text{IV}} = (X_{1j}' P_{G_j} X_{1j})^{-1} X_{1j}' P_{G_j} Y_{1j} \tag{22}$$

and $\hat{\theta}_{j2\text{IV}} = 0$, respectively. These are obtained by multiplying the matrix P_{G_j} to both sides of equation (21) and running the principal components OLS regression.

Because $X_{1j}' P_{G_j} X_{2j} = 0$ because of the orthogonality of eigenvectors, equation (22) can be rewritten as

$$\hat{\theta}_{j1\text{IV}} = S_{j1}'\delta + (X_{1j}' P_{G_j} X_{1j})^{-1} X_{1j}' P_{G_j} U_j. \tag{23}$$

But the second term on the right-hand side of equation (23) is asymptotically negligible, and hence we have

$$\begin{bmatrix} \hat{\theta}_{11\text{IV}} \\ \hat{\theta}_{21\text{IV}} \end{bmatrix} \cong \begin{bmatrix} S_{11}' \\ S_{21}' \end{bmatrix} \delta = \Gamma\delta.$$

Applying the principle of minimum distance estimation as in Section 3 gives the PCIV estimator of the vector δ as[11]

$$\hat{\delta}_{\text{PCIV}} = (\Gamma' \Theta_\delta^{-1} \Gamma)^{-1} \Gamma' \Theta_\delta^{-1} \begin{bmatrix} \hat{\theta}_{11\text{IV}} \\ \hat{\theta}_{21\text{IV}} \end{bmatrix}, \tag{24}$$

where $\Theta_\delta = \begin{bmatrix} (X_{11}' P_{G_1} X_{11})^{-1} & 0 \\ 0 & (X_{12}' P_{G_2} X_{12})^{-1} \end{bmatrix}$.

The PCIV estimator for the regression equation (10) is defined in the same way as for the regression equation (1). Write model (10) under $i \in \mathcal{G}_j$ as

$$\begin{aligned} Y_{1j} &= \overline{W}_j \phi + U_j \\ &= \overline{W}_j \bar{S}_j \bar{S}_j' \phi + U_j \\ &= \overline{W}_j \bar{S}_{j1} \tau_{j1} + \overline{W}_j \bar{S}_{j2} \tau_{j2} + U_j \\ &= \overline{X}_{1j} \tau_{j1} + \overline{X}_{2j} \tau_{j2} + U_j, \end{aligned}$$

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

where $\bar{S}_j = \begin{bmatrix} k_1+k_2 & n \\ \bar{S}_{j1} & \bar{S}_{j2} \end{bmatrix}$ is a matrix whose columns are orthonormal eigenvectors of $\overline{W}_j' P_{\bar{G}_j} \overline{W}_j$, $\bar{S}_{j1}$ corresponds to the $k_1 + k_2$ largest eigenvalues of $\overline{W}_j' P_{\bar{G}_j} \overline{W}_j$, $\tau_{jm} = \bar{S}_{jm}' \phi$ $(m = 1,2)$ and $\overline{X}_{1j} = \overline{W}_j \bar{S}_{jm}$.

By using part (b) of Lemma 1, the PCIV estimator of τ_{j1} is defined as

$$\bar{\tau}_{j1\mathrm{IV}} = (\overline{X}_{1j}' P_{\bar{G}_j} \overline{X}_{1j})^{-1} \overline{X}_{1j}' P_{\bar{G}_j} Y_{1j}.$$

But because

$$\bar{\tau}_{j1\mathrm{IV}} = \bar{S}_{j1}' \phi + (\overline{X}_{1j}' P_{\bar{G}_j} \overline{X}_{1j})^{-1} \overline{X}_{1j}' P_{\bar{G}_j} U_j, \tag{25}$$

and the second term on the right-hand side of equation (25) is negligible, the PCIV estimator of ϕ is defined as

$$\bar{\phi}_{\mathrm{PCIV}} = (\bar{\Gamma}' \Theta_\phi^{-1} \bar{\Gamma})^{-1} \bar{\Gamma}' \Theta_\phi^{-1} \begin{bmatrix} \bar{\tau}_{11\mathrm{IV}} \\ \bar{\tau}_{21\mathrm{IV}} \end{bmatrix}, \tag{26}$$

where $\bar{\Gamma} = \begin{bmatrix} \bar{S}_{11}' \\ \bar{S}_{21}' \end{bmatrix}$ and $\Theta_\phi = \begin{bmatrix} (\overline{X}_{11}' P_{\bar{G}_1} \overline{X}_{11})^{-1} & 0 \\ 0 & (\overline{X}_{12}' P_{\bar{G}_2} \overline{X}_{12})^{-1} \end{bmatrix}$.

4.3. Asymptotic Distribution Theory

This section studies asymptotic distributions of the PCIV estimators introduced in the last section. To derive the asymptotic distributions of the PCIV estimators $\hat{\delta}_{\mathrm{PCIV}}$ and $\bar{\phi}_{\mathrm{PCIV}}$, the following lemma will be used.

LEMMA 2.

(a) When $N_j \to \infty$, under Assumptions 1–4 and 8,

$$\sqrt{N_j}(\hat{\theta}_{j1\mathrm{IV}} - S_{j1}' \delta) \xrightarrow{d} N\left(0, \sigma_{uu} \mathrm{diag}\left(\frac{1}{\xi_{j1}}, \ldots, \frac{1}{\xi_{jk_1}}\right)\right) = N(0, \sigma_{uu} V_{\delta j}), \text{ say,}$$

where $\xi_{j1}, \ldots, \xi_{jk_1}$ are k_1 largest eigenvalues of the matrix $M_{W_j G_j} M_{G_j G_j}^{-1} M_{G_j W_j} = \mathrm{plim}_{N_j \to \infty} (1/N_j) W_j' P_{G_j} W_j$.

(b) When $N_j \to \infty$, under Assumptions 1, 3, 5, 6, and 8,

$$\sqrt{N_j}(\bar{\tau}_{j1\mathrm{IV}} - \bar{S}_{j1}' \phi) \xrightarrow{d} N\left(0, \sigma_{uu} \mathrm{diag}\left(\frac{1}{\bar{\xi}_{j1}}, \ldots, \frac{1}{\bar{\xi}_{j(k_1+k_2)}}\right)\right) = N(0, \sigma_{uu} \overline{V}_{\phi j}), \text{ say,}$$

where $\bar{\xi}_{j1}, \ldots, \bar{\xi}_{j(k_1+k_2)}$ are $k_1 + k_2$ largest eigenvalues of the matrix $M_{\overline{W}_j \bar{G}_j} M_{\bar{G}_j \bar{G}_j}^{-1} M_{\bar{G}_j \overline{W}_j} = \mathrm{plim}_{N_j \to \infty} (1/N_j) \overline{W}_j' P_{\bar{G}_j} \overline{W}_j$.

Remark. This lemma shows that $\hat{\theta}_{j1\mathrm{IV}}$ and $\bar{\tau}_{j1\mathrm{IV}}$ estimate particular linear combinations of the structural parameter vectors δ and ϕ consistently and that their

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

asymptotic distributions are normal. Under the special case $S_{j2}'\delta = 0$, the parameter vector δ can be estimated consistently by $S_{j1}\hat{\theta}_{j1\mathrm{IV}}$, because

$$S_{j1}\hat{\theta}_{j1\mathrm{IV}} - S_{j1}S_{j1}'\delta = S_{j1}\hat{\theta}_{j1\mathrm{IV}} - (I - S_{j2}S_{j2}')\delta = S_{j1}\hat{\theta}_{j1\mathrm{IV}} - \delta.$$

No doubt, such a special case would rarely occur in practice. Likewise, when $\bar{S}_{j2}'\phi = 0$, the parameter vector ϕ can be estimated consistently by $\bar{S}_{j1}\bar{\tau}_{j1\mathrm{IV}}$.

Using Lemma 2 and formulae (24) and (26), the asymptotic distributions of the PCIV estimators for δ and ϕ can be obtained as follows.

THEOREM 2.

(a) Under Assumptions 1–4 and 8, as $N, N_j \to \infty$ $(j = 1,2)$,

$$\sqrt{N}(\hat{\delta}_{\mathrm{PCIV}} - \delta) \xrightarrow{d} N(0, \sigma_{uu}(\Psi'\Delta_\delta\Psi)^{-1}),$$

where Ψ *is the probability limit of* Γ, $\Delta_\delta = \begin{bmatrix} d_1 V_{\delta 1}^{-1} & 0 \\ 0 & d_2 V_{\delta 2}^{-1} \end{bmatrix}$, *and* d_j *and* $V_{\delta j}$ $(j = 1,2)$ *are defined in Assumption 8 and Lemma 2, respectively.*

(b) Under Assumptions 1, 3, 5, 6, and 8, as $N, N_j \to \infty$ $(j = 1,2)$,

$$\sqrt{N}(\bar{\phi}_{\mathrm{PCIV}} - \phi) \xrightarrow{d} N(0, \sigma_{uu}(\bar{\Psi}'\bar{\Delta}_\phi\bar{\Psi})^{-1}),$$

where $\bar{\Psi}$ *is the probability limit of* $\bar{\Gamma}$ *and* $\bar{\Delta}_\phi = \begin{bmatrix} d_1 \bar{V}_{\phi 1}^{-1} & 0 \\ 0 & d_2 \bar{V}_{\phi 2}^{-1} \end{bmatrix}$ *and* d_j *and* $\bar{V}_{\phi j}$ $(j = 1,2)$ *are defined in Assumption 8 and Lemma 2, respectively.*

Remarks.

(a) A simple calculation yields

$$M_{W_jG_j}M_{G_jG_j}^{-1}M_{G_jW_j} = \begin{bmatrix} \Pi_{1j}'M_{z_{1j}z_{1j}}\Pi_{1j} & \Pi_{1j}'M_{z_{1j}z_{1j}} \\ M_{z_{1j}z_{1j}}\Pi_{1j} & M_{z_{1j}z_{1j}} \end{bmatrix}. \tag{27}$$

Accordingly, the asymptotic variance-covariance matrix for $\hat{\delta}_{\mathrm{PCIV}}$ depends only on $M_{z_{1j}z_{1j}}$ and Π_{1j} $(j = 1,2)$. In the same manner, the asymptotic variance-covariance matrix for $\bar{\phi}_{\mathrm{PCIV}}$ depends only on $M_{z_{1j}z_{1j}}, M_{z_{1j}z_{2j}}, M_{z_{2j}z_{2j}}, \Pi_{1j}$, and Π_2, because

$$M_{\bar{W}_j\bar{G}_j}M_{\bar{G}_j\bar{G}_j}^{-1}M_{\bar{G}_j\bar{W}_j} = \begin{bmatrix} \Pi^{j\prime}M_{\bar{z}_j\bar{z}_j}\Pi^j & \Pi^{j\prime}M_{\bar{z}_j\bar{z}_j} \\ M_{\bar{z}_j\bar{z}_j}\Pi^j & M_{\bar{z}_j\bar{z}_j} \end{bmatrix},$$

where $\Pi^j = \begin{bmatrix} \Pi_{1j} \\ \Pi_2 \end{bmatrix}$ and $M_{\bar{z}_j\bar{z}_j} = \begin{bmatrix} M_{z_{1j}z_{1j}} & M_{z_{1j}z_{2j}} \\ M_{z_{2j}z_{1j}} & M_{z_{2j}z_{2j}} \end{bmatrix}$. It is interesting to observe that such limiting moments as $M_{h_jh_j}, M_{h_jz_{1j}}, M_{f_jf_j}, M_{f_jz_{1j}}$, and $M_{f_jz_{2j}}$ do not affect the asymptotic variance-covariance matrices. This implies that the asymptotic efficiency of the PCIV estimators is not affected by the choice of instrument matrices h_i and f_i as long as they are uncorrelated with the disturbance term u_i. This counter-intuitive phenomenon occurs because the PCIV regression methods use only some largest eigenvalues of the projected sample moment matrix.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

(b) When $\Pi_{11} = \Pi_{12}$, we would have $M_{W_1G_1}M_{G_1G_1}^{-1}M_{G_1W_1} \simeq M_{W_2G_2}M_{G_2G_2}^{-1}M_{G_2W_2}$ because it is reasonable to expect that $M_{z_{11}z_{11}} \simeq M_{z_{12}z_{12}}$. Therefore, the matrix $\Psi'\Psi$ is nearly singular, and the variance-covariance matrix in part (a) of Theorem 2 is not well defined. This also holds for part (b) of Theorem 2 under $\Pi_{11} = \Pi_{12}$.

(c) Now, assume the polar opposite case of Assumption 4 for the regression equation (1), that is, the case of $\Pi_{11} = \Pi_{12} = 0$. Then, because $S'_{j11} = 0$ where $S'_{j1} = [S'^{n}_{j11} \quad S'^{k_1}_{j12}]$, it is deduced from equation (23)

$$\hat{\theta}_{j1\mathrm{IV}} \xrightarrow{p} \mathrm{plim}_{N_j\to\infty} S'_{j12}\gamma.$$

This implies that $\hat{\theta}_{j1\mathrm{IV}}$ does not contain any information on the parameter vector β and that β cannot be estimated consistently. But the parameter vector γ can be estimated consistently[12] once we change the matrix Γ such that $\Gamma = \begin{bmatrix} S'_{112} \\ S'_{212} \end{bmatrix}$. The asymptotic distribution of the PCIV estimator for γ can be recovered from Theorem 2. Intuitively speaking, the instruments we have chosen have no correlation with the regressor y_{2i} under $\Pi_{11} = \Pi_{12} = 0$, and hence they cannot be valid instruments for y_{2i}, which results in inconsistent estimation for β. In a more general case where $[\Pi'_{11}, \Pi'_{12}]'$ is of deficient rank, we may also find by rotating the matrix $[\Pi'_{11}, \Pi'_{12}]'$ as in Phillips (1989) and extending the analysis given earlier that β cannot be estimated consistently but γ can be.[13]

(d) When the polar opposite case of Assumption 4 for equation (10) holds (i.e., $\Pi_{11} = \Pi_{12} = 0$) and when $\Pi_2 = 0$, the first n columns of $\bar{S}'_{j11}$ are null vectors, and it follows as for equation (1) that β cannot be estimated consistently but that γ and ρ can be. We may show that this also holds when $[\Pi'_{11}, \Pi'_{12}]'$ is of deficient column rank and $\Pi_2 = 0$. In another polar case where $\Pi_{11} = \Pi_{12} = 0$ and $\mathrm{rank}(\Pi_2) = n$, $M_{\bar{W}_1\bar{G}_1}M_{\bar{G}_1\bar{G}_1}^{-1}M_{\bar{G}_1\bar{W}_1} \simeq M_{\bar{W}_2\bar{G}_2}M_{\bar{G}_2\bar{G}_2}^{-1}M_{\bar{G}_2\bar{W}_2}$ because it is reasonable to expect that $M_{\bar{z}_1\bar{z}_1} \simeq M_{\bar{z}_2\bar{z}_2}$. Therefore, $\bar{\Psi}'\bar{\Psi}$ becomes nearly singular, and the asymptotic variance-covariance matrix of $\sqrt{N}(\bar{\phi}_{PCIV} - \phi)$ is not well defined. This can be shown to hold in a more general case where $[\Pi'_{11}, \Pi'_{12}]'$ is of deficient rank and $\mathrm{rank}(\Pi_2) = n$.

(e) It would be interesting to specialize on the case $n = k_1 = 1$ for the regression equation (1), because this simple case would explicitly reveal the nonlinear influence of Π_{1j} and $M_{z_{1j}z_{1j}}$ on the asymptotic variances of $\hat{\beta}_{\mathrm{PCIV}}$ and $\hat{\gamma}_{\mathrm{PCIV}}$. Using relation (27), the nonzero eigenvalue of $M_{W_jG_j}M_{G_jG_j}^{-1}M_{G_jW_j}$ and the corresponding eigenvector are obtained as

$$\xi_{j1} = M_{z_{1j}z_{1j}} + \Pi_{1j}^2 M_{z_{1j}z_{1j}}, \qquad \mathrm{plim}_{N_j\to\infty} S_{j1} = \begin{bmatrix} \dfrac{\Pi_{1j}}{\sqrt{\Pi_{1j}^2 + 1}} \\ \dfrac{1}{\sqrt{\Pi_{1j}^2 + 1}} \end{bmatrix}.$$

Using these and part (a) of Theorem 2, we find that the asymptotic variance-covariance matrix of $\hat{\delta}_{\mathrm{PCIV}}$ is equivalent to that of the MD estimator given in (19). Thus, the same interpretations as in Section 3.2 can be given here.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

5. MONTE CARLO SIMULATION

This section reports the empirical biases, standard deviations, and root mean squared errors of the MD and PCIV estimators. Data for our experiments were generated according to

$$y_{1i} = \beta y_{2i} + \gamma z_{1i} + u_i, \qquad \beta = \gamma = 1 \text{ (for equation (1))};$$

$$y_{1i} = \beta y_{2i} + \gamma z_{1i} + \rho z_{2i} + u_i, \qquad \beta = \gamma = \rho = 1 \text{ (for equation (10))};$$

$$y_{2i} = \Pi_{11} z_{1i} + v_{2i} \qquad (i = 1,\ldots,N_1); \Pi_{11} = 1;$$

$$y_{2i} = \Pi_{12} z_{1i} + v_{2i} \qquad (i = N_1 + 1,\ldots,N_2);$$

$$z_{1i} \sim iid\,(\chi_1^2 - 1); z_{2i} \sim iid\,(\chi_1^2 - 1);$$

$$\begin{bmatrix} u_i \\ v_{2i} \end{bmatrix} \sim iid\, N\left(0, \begin{bmatrix} 1 & \sigma_{uv} \\ \sigma_{uv} & 1 \end{bmatrix}\right),$$

where z_{1i} and z_{2i} were fixed throughout iterations and the number of iterations was set at 20,000. The PCIV estimator used $[z_{1i}^2, z_{1i}]$ as instruments for the seemingly unidentified structural equation under the rank condition failure and $[z_{1i}^2, z_{1i}, z_{2i}]$ for the seemingly unidentified structural equation under the order condition failure.

First, we considered the case where the structural change points are known and utilized correctly in estimation. The experimental results for this case are reported in Table 1. Part I of Table 1 considers the seemingly unidentified structural equation under the rank condition failure (equation (1)) and part II that under the order condition failure (equation (10)). The experimental results in Table 1 can be summarized as follows.

- Part I shows that the PCIV estimator tends to have less standard deviation and root mean squared error than the MD estimator at $(N_1, N_2) = (50,50)$. But at $(N_1, N_2) = (100,100)$, the converse is true. Regarding the bias, it is not certain which estimator performs better at $(N_1, N_2) = (50,50)$. But at $(N_1, N_2) = (100,100)$, the MD estimator tends to be less biased than the PCIV estimator.
- According to Part II, the PCIV estimator is less biased than the MD estimator excepting the case $(N_1, N_2) = (50,50), \sigma_{uv} = .5$ and $\Pi_{12}/\Pi_{11} = .25$. Furthermore, the PCIV estimator has less standard deviation and root mean squared error than the MD estimator in all cases.
- As the value of Π_{12}/Π_{11} increases, the absolute value of bias, standard deviations, and root mean squared errors tend to decrease though there are a few exceptions.
- The effect of σ_{uv} on bias, standard deviations, and root mean squared errors appears to be uncertain. But as the value of σ_{uv} increases, the absolute value of bias tends to increase, though there are a few exceptions at $(N_1, N_2) = (50,50)$.

In practice, known structural change points may not be the true ones. Table 2 reports simulation results that use misspecified structural change points. The misspecified structural change points were selected randomly from intervals [41,50] and [52,61] at $(N_1, N_2) = (50,50)$ and from intervals [81,100] and

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

Table 1. Biases, standard deviations, and root mean square errors with known change points

				β			γ		
	N_1, N_2	σ_{uv}	Π_{12}/Π_{11}	Bias	Std	RMSE	Bias	Std	RMSE
				Part I. The case of rank condition failure					
MD	50, 50	.5	.25	1.94	228	228	−2.24	266	266
			.50	−.14	3.15	3.16	.16	3.48	3.48
		.75	.25	.60	46.9	46.9	−.62	50.1	50.1
			.50	−.07	16.1	16.1	.08	18.0	18.0
	100, 100	.5	.25	−.13	6.95	6.95	.16	7.75	7.75
			.50	−.03	.25	.26	.04	.33	.33
		.75	.25	−.30	19.7	19.7	.36	23.9	23.9
			.50	−.04	.27	.27	.06	.35	.36
PCIV	50, 50	.5	.25	−.11	76.4	76.4	.19	84.8	84.8
			.50	−.36	29.5	29.6	.41	32.2	32.2
		.75	.25	.06	30.3	30.3	−.10	37.0	37.0
			.50	−.12	4.49	4.49	.15	4.83	4.84
	100, 100	.5	.25	−.37	27.4	27.4	.41	31.3	31.3
			.50	−.03	.26	.26	.04	.33	.33
		.75	.25	−.58	47.1	47.1	.69	56.5	55.9
			.50	−.05	.27	.28	.06	.36	.36

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

	N_1, N_2	σ_{uv}	Π_{12}/Π_{11}	β			γ			ρ		
				Bias	Std	RMSE	Bias	Std	RMSE	Bias	Std	RMSE
Part II. The case of order condition failure												
MD	50, 50	.5	.25	−0.04	21.2	21.2	.05	22.1	22.1	.02	1.60	1.60
			.50	.08	25.7	25.7	−.10	30.4	30.4	.00	.31	.31
		.75	.25	.96	132	132	−.93	136	136	.00	2.58	2.58
			.50	−.22	8.96	8.96	.26	10.5	10.5	.00	.29	.29
	100, 100	.5	.25	−.13	7.94	7.94	.15	9.13	9.13	.00	.27	.27
			.50	−.03	.26	.26	.04	.33	.34	.00	.05	.05
		.75	.25	−.17	10.8	10.8	.20	12.7	12.7	.00	.50	.50
			.50	−.04	.27	.28	.06	.36	.36	.00	.05	.05
PCIV	50, 50	.5	.25	.20	1.29	1.31	−.21	1.39	1.40	.00	.14	.14
			.50	.03	1.00	1.00	−.03	1.10	1.10	.00	.11	.11
		.75	.25	.31	1.57	1.60	−.33	1.66	1.69	.00	.15	.15
			.50	.03	.55	.56	−.03	.64	.64	.00	.09	.09
	100, 100	.5	.25	.04	.92	.92	−.04	1.04	1.04	.00	.06	.06
			.50	.00	.24	.24	.00	.31	.31	.00	.05	.05
		.75	.25	.06	1.41	1.41	−.06	1.71	1.71	.00	.08	.08
			.50	.00	.25	.25	.00	.32	.32	.00	.05	.05

Note: Calculations are based on 20,000 iterations; $\Pi_{11} = 1$.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

TABLE 2. Biases, standard deviations, and root mean square errors with misspecified change points

				β			γ		
	N_1, N_2	σ_{uv}	Π_{12}/Π_{11}	Bias	Std	RMSE	Bias	Std	RMSE
				Part I. The case of rank condition failure					
MD	50, 50	.5	.25	−.31	53.1	53.1	.37	58.0	58.0
			.50	−.18	16.1	16.1	.22	18.9	18.9
		.75	.25	−.24	52.3	52.3	.30	55.3	55.3
			.50	−.36	67.4	67.4	.36	71.8	71.8
	100, 100	.5	.25	−.16	11.9	11.9	.18	13.3	13.3
			.50	−.05	.65	.65	.06	.86	.86
		.75	.25	.91	201	201	−.98	223	223
			.50	−.06	1.24	1.25	.07	1.58	1.58
PCIV	50, 50	.5	.25	−1.20	216	216	1.26	226	226
			.50	.46	64.3	64.3	−.53	73.2	73.2
		.75	.25	−1.18	181	181	1.19	185	185
			.50	−.56	52.3	52.3	.69	64.7	64.7
	100, 100	.5	.25	.13	33.0	33.0	−.16	38.3	38.3
			.50	−.05	.42	.42	.06	.55	.55
		.75	.25	.18	49.7	49.7	−.18	55.0	55.0
			.50	−.09	3.51	3.51	.12	4.23	4.24

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

				β			γ			ρ		
	N_1, N_2	σ_{uv}	Π_{12}/Π_{11}	Bias	Std	RMSE	Bias	Std	RMSE	Bias	Std	RMSE
Part II. The case of order condition failure												
MD	50, 50	.5	.25	2.92	295	295	−3.01	298	298	.04	6.31	6.31
			.50	−.10	38.7	38.7	.09	41.7	41.7	.01	1.60	1.60
		.75	.25	.03	91.0	91.0	.01	95.0	95.0	−.06	5.72	5.72
			.50	.06	60.9	60.9	−.05	72.0	72.0	.02	1.39	1.39
	100, 100	.5	.25	−.07	30.6	30.6	.09	35.1	35.1	.00	.99	.99
			.50	−.04	.45	.45	.06	.58	.59	.00	.05	.05
		.75	.25	2.84	308	308	−3.25	354	354	.05	4.72	4.72
			.50	−.07	.68	.69	.09	.85	.86	.00	.05	.05
PCIV	50, 50	.5	.25	.32	1.64	1.67	−.34	1.74	1.77	.00	.14	.14
			.50	.11	.92	.93	−.12	1.06	1.07	.00	.12	.12
		.75	.25	.51	7.41	7.43	−.54	7.66	7.67	.00	.40	.40
			.50	.15	1.22	1.23	−.17	1.41	1.42	.00	.11	.11
	100, 100	.5	.25	.08	.94	.94	−.09	1.98	1.09	.00	.06	.06
			.50	.00	.30	.30	.00	.39	.39	.00	.05	.05
		.75	.25	.12	1.02	1.03	−.13	1.16	1.17	.00	.06	.06
			.50	.00	.35	.35	.00	.46	.46	.00	.05	.05

Notes: Calculations are based on 20,000 iterations; $\Pi_{11} = 1$. The misspecified structural change points were selected randomly from intervals [41,50] and [52,61] at $N_1 = N_2 = 50$ and from intervals [81,100] and [102,121] at $N_1 = N_2 = 100$.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

TABLE 3. Biases, standard deviations, and root mean square errors with estimated change points

				β			γ		
	N_1, N_2	σ_{uv}	Π_{12}/Π_{11}	Bias	Std	RMSE	Bias	Std	RMSE
				Part I. The case of rank condition failure					
MD	50, 50	.5	.25	.28	.68	.73	−.30	.73	.79
			.50	.07	.43	.43	−.08	.49	.50
		.75	.25	.42	.64	.77	−.45	.69	.82
			.50	.11	.44	.45	−.12	.50	.52
	100, 100	.5	.25	.07	.57	.57	−.07	.65	.66
			.50	.01	.23	.23	−.02	.30	.30
		.75	.25	.10	.61	.62	−.11	.70	.71
			.50	.02	.24	.24	−.03	.31	.31
PCIV	50, 50	.5	.25	.24	10.6	10.6	−.26	12.2	12.2
			.50	−.06	8.08	8.08	.07	8.49	8.49
		.75	.25	.39	15.1	15.1	−.41	16.4	16.4
			.50	−.00	3.64	3.64	.01	4.2	4.2
	100, 100	.5	.25	.01	1.99	1.99	−.01	2.34	2.34
			.50	.00	.24	.24	.00	.31	.31
		.75	.25	.14	8.59	8.59	−.16	9.86	9.86
			.50	.00	.25	.25	.00	.32	.32

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

	N_1, N_2	σ_{uv}	Π_{12}/Π_{11}	β			γ			ρ		
				Bias	Std	RMSE	Bias	Std	RMSE	Bias	Std	RMSE
Part II. The case of order condition failure												
MD	50, 50	.5	.25	.22	4.24	4.24	−.23	4.22	4.22	.00	.45	.45
			.50	.07	.49	.49	−.08	.55	.56	.00	.09	.09
		.75	.25	.45	3.80	3.83	−.47	3.59	3.63	.00	.25	.25
			.50	.10	1.22	1.22	−.11	1.38	1.39	.00	.18	.18
	100, 100	.5	.25	.06	.58	.59	−.07	.67	.67	.00	.05	.05
			.50	.01	.23	.23	−.02	.30	.30	.00	.05	.05
		.75	.25	.09	.73	.73	−.11	.84	.85	.00	.05	.05
			.50	.02	.24	.24	−.03	.31	.31	.00	.05	.05
PCIV	50, 50	.5	.25	.33	.85	.91	−.35	.94	1.00	.00	.01	.01
			.50	.12	.53	.54	−.13	.61	.63	.00	.09	.09
		.75	.25	.54	7.33	7.35	−.57	7.57	7.59	.00	.39	.39
			.50	.18	.63	.65	−.20	.73	.75	.00	.09	.09
	100, 100	.5	.25	.12	.53	.54	−.13	.61	.62	.00	.05	.05
			.50	.03	.23	.23	−.03	.30	.30	.00	.05	.05
		.75	.25	.19	.54	.57	−.21	.62	.66	.00	.05	.05
			.50	.04	.23	.23	−.05	.30	.30	.00	.04	.04

Notes: Calculations are based on 20,000 iterations; $\Pi_{11} = 1$. The structural change point was estimated by the least squares method, assuming there is only one structural change.

[102,121] at $(N_1, N_2) = (100,100)$. Note that the true change points are 51 and 101 at $(N_1, N_2) = (50,50)$ and $(N_1, N_2) = (100,100)$, respectively. The simulation results can be summarized as follows.

- As well expected, the estimators perform worse than those using correct change points, though there is an exception for the MD estimator at $(N_1, N_2) = (50,50)$.
- Part I shows that the MD estimator tends to perform better than the PCIV estimator at $(N_1, N_2) = (50,50)$ in terms of both bias and standard deviation. At $(N_1, N_2) = (100,100)$, though, it seems difficult to decide which estimator performs better.
- According to Part II, the PCIV estimator performs uniformly better than the MD estimator in terms of standard deviation and root mean squared errors. The PCIV estimator also tends to be less biased.

The structural change points may not be known and need to be estimated in practice. Table 3 reports simulation results that use estimated structural change points. Assuming that there is only one structural change, the structural change points were estimated by the least squares method.[14] The results can be summarized as follows.

- Tables 1 and 3 show that bias, standard deviations, and root mean squared errors tend to improve when estimated structural change points are used. For the MD estimator, this may be explained by the fact that the objective function for the least squares estimation of structural change point is related to the variance-covariance estimator of the MD estimator. For the PCIV estimator, it is less obvious why the estimated structural change points bring better results.
- Comparing the MD and PCIV estimators, the MD estimator tends to have less standard deviation and root mean squared error than the PCIV estimator in Part I, but the converse is true in Part II. Regarding the bias, the PCIV estimator tends to have less bias in Part I, but the converse holds in Part II.
- When $(N_1, N_2) = (100,100)$ and $\Pi_{12}/\Pi_{11} = .5$, the estimated structural change points are close to the true ones, and therefore the performance of the MD and PCIV estimator in Table 3 is similar to that in Table 1.

6. AN EMPIRICAL EXAMPLE

This section studies the Japanese export-GDP relation by using the methods proposed in previous sections and the Japanese deseasonalized quarterly data covering the period 1960:Q2–1998:Q3. Most previous time series studies investigating the export-GDP relation use the VAR causality tests (e.g., Shan and Sun, 1999; Marin, 1992; Ghartey, 1993). However, the VAR causality tests ignore contemporaneous relations which seem quite important in our empirical studies of Japanese exports and GDP, as will be seen shortly.

Instead of the VAR causality tests, we consider using the simultaneous equations system

$$GDPR_i = \beta EXPR_i + \gamma' \begin{pmatrix} 1 \\ EXPR_{i-1} \end{pmatrix} + \rho GDPR_{i-1} + u_i, \tag{28}$$

$$EXPR_i = \Pi_1' \begin{pmatrix} 1 \\ EXPR_{i-1} \end{pmatrix} + v_{2i}, \tag{29}$$

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

where $GDPR_i$ denotes the GDP growth rate and $EXPR_i$ the export growth rate.[15] Equation (28) signifies the structural equation and equation (29) the reduced form equation for $EXPR_i$. Equations (28) and (29) do not satisfy the order condition for identification and hence cannot be estimated by conventional methods. The Japanese quarterly data seem to support the AR(1) specification in equation (29). Notably, the lagged GDP growth rates do not seem to affect the current export growth rate. In equation (28), it may be more appropriate to include the variable $GDPR_{i-2}$.[16] But the results using the extra variable indicate that the coefficients for the export variables are not significant, which is counterintuitive. So we will focus only on equation (28).

The least squares methods for detecting structural change points (cf. Bai and Perron, 1998) find that there is a structural change in the intercept term of equation (29) at the period 1982:Q2. Moreover, the sup-Wald test of Andrews (1993) also indicates[17] that there is one significant structural change in the intercept term of equation (29). Indeed, the average quarterly export growth rate before 1982:Q2 was 3.6%, whereas that since 1982:Q2 was only 0.7%. Thus, we conclude that the Japanese export growth rate underwent a structural change during the sample period.

Now, assuming that the structural parameters were constant during the sample period,[18] we estimated the coefficients of equation (28) by the MD and PCIV methods. The estimation results are reported in Table 4. As nonlinear instruments, $EXPR^2_{i-1}, EXPR^3_{i-1}$, and $(GDPR^2_{i-1}, EXPR^2_{i-1})$ were used for the PCIV estimator. The structural change point was set at 1982:Q2.

The results in Table 4 indicate that all the regressors are significant at conventional significance levels. Because the averages of $GDPR_i$ and $GDPR_{i-1}$ and those of $EXPR_i$ and $EXPR_{i-1}$ are approximately the same, respectively, and because the intercept term is statistically insignificant, the estimation results from the MD estimation yield a relation

$$Avg(GDPR) \cong 0.583Avg(EXPR) - 0.223Avg(EXPR) + 0.371Avg(GDPR)$$

Table 4. PCIV estimation results for the Japanese export-GDP relation

Estimator	$EXPR$	$EXPR(-1)$	$GDPR(-1)$	$Avg(GDPR)/Avg(EXPR)$
MD	.583 (2.11*)	−.223 (−1.84*)	.371 (2.76**)	.57
PCIV1	.636 (3.00**)	−.246 (−2.39*)	.362 (2.74**)	.61
PCIV2	.630 (3.04**)	−.244 (−2.39*)	.365 (2.81**)	.61
PCIV3	.407 (3.48**)	−.155 (−2.48*)	.453 (4.48**)	.46

Notes: The numbers in parentheses are *t*-ratios. PCIV1, PCIV2, and PCIV3 use $EXPR^2_{i-1}, EXPR^3_{i-1}$, and $(EXPR^2_{i-1}, GDPR^2_{i-1})$ as nonlinear instruments, respectively.
**: significant at the 1% level. *: significant at the 5% level.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

TABLE 5. OLS estimation results for the Japanese export-GDP relation

EXPR	*EXPR*(−1)	*GDPR*(−1)	*Avg*(*GDPR*)/*Avg*(*EXPR*)
.722 (3.67)	−.113 (−.54)	.404 (5.40)	1.02

Note: The numbers in parentheses are *t*-ratios.

or

$$Avg(GDPR) \cong \frac{0.583 - 0.223}{1 - 0.371} Avg(EXPR) = 0.57 Avg(EXPR), \quad \textbf{(30)}$$

where $Avg(A)$ denotes the sample mean of A_i. We may infer from this relation that the 1% increase of export growth rate had the effect of increasing GDP by 0.57% on the average. Using the PCIV estimation results, we obtained similar numbers as reported in Table 4. In light of the fact that the average ratio of export to GDP was 11.2% during the sample period in Japan, relation (30) and the numbers in the last column of Table 4 seem to suggest that export had an accelerating multiplier effect on the GDP growth in Japan during the sample period.

The OLS estimation for equation (28) is subject to the problems of endogeneity and multicollinearity. To illustrate how serious these can be and to compare the results with those in Table 4, the OLS estimation was performed for equation (28). The results are reported in Table 5. Using the same arguments as before, we may interpret the results such that the 1% increase of export growth rate brings the 1.022% growth of GDP. This number seems to be too large relative to previous cross-sectional regression results in the literature (e.g., Lee and Cole, 1994). This must be due to the problems associated with estimating equation (28) by OLS.

7. SUMMARY AND FURTHER REMARKS

We have studied structural equations that do not satisfy the rank and/or order condition(s) for identification but still are identifiable. To estimate the seemingly unidentified structural equations, we proposed the MD and PCIV estimators. The asymptotic distributions of these estimators were reported. In a special case, these estimators were shown to have the same asymptotic distribution. But the asymptotic equivalence of these estimators in general could not be derived in this paper. Simulation results indicated that the MD and PCIV estimators were complementary to each other. But it was also found that the PCIV estimator tended to perform better in the case of order condition failure. The estimation methods proposed in this paper were applied to the Japanese export

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

and GDP data. The results revealed that export had an accelerating multiplier effect on the GDP growth in Japan during the sample period 1960:Q2–1998:Q3.

There are a few topics that deserve further investigation but were not studied in this paper because of space limitations. First, the new identification conditions suggested in this paper are the rank conditions implying the presence of structural changes in the reduced form equation (Assumptions 4 and 6) and the absence of structural changes in the structural equation. The former condition can be tested by existing methods such as those of Gill and Lewbel (1992) and Cragg and Donald (1996), though these require modifications when structural change points are not known. But it seems that the latter condition cannot be tested by existing methods, and therefore new methods need to be developed. Second, this paper assumes that the structural change point is known. But estimating structural coefficients and unknown structural change points simultaneously would be a useful topic for future research. Third, this paper considers the case of identification failure in the conventional sense, but studying how structural changes affect the estimator properties in the case of weak instruments as in Staiger and Stock (1997) would be an interesting future research topic. Fourth, in light of prevalent evidence that many economic time series are nonstationary, it would be meaningful to study the seemingly unidentified structural equations under unit root nonstationarity.

NOTES

1. Hausman (1983), Hsiao (1983), Mariano (1982), and Phillips (1983) survey earlier contributions in this research area.

2. This estimator was kindly suggested by a referee.

3. For simplicity, we consider only two data regimes (or one structural change), but extension of the results in this paper to the case of several data regimes (or multiple structural changes) is straightforward.

4. Finite sample and asymptotic distribution theory of IV estimators under the assumption $\text{rank}(\Pi_2) < n$, which covers Assumption 2 as a special case, is studied in Phillips (1989) and Choi and Phillips (1992). In addition, Staiger and Stock (1997) study the finite sample and asymptotic distribution theory of IV estimators when $\Pi_2 = C/\sqrt{T}$ where C is a constant matrix.

5. Testing the null hypothesis $H_0: \rho = 0$ has attracted much attention from econometricians, and the tests for this null have been called overidentification tests (cf. Koopmans and Hood, 1953; Anderson and Kunitomo, 1992; Cragg and Donald, 1993).

6. See, for example, Newey and McFadden (1994).

7. See Judge and Bock (1983); Judge, Griffiths, Hill, Lütkepohl, and Lee (1985) and references therein.

8. When $n = k_1 = 1$, examples of instruments are z_{1i}^2, z_{1i}^4, and so on. Noninteger moments of z_{1i} can also be used as instruments when z_{1i} is nonnegative. When $z_{1i} = [1, x_{1i}']'$, that is, when z_{1i} includes a constant term, we use $H(x_{1i})$ as instruments.

9. However, using the nonlinear functions of z_{1i} is not the main key to the estimability of unidentified structural equations. When other instruments are available and uncorrelated with the disturbance term u_i, they can also be used as instruments.

10. See the recent work by Lewbel (1997) and Dagenais and Dagenais (1995) and references therein.

11. Both $\hat{\theta}_{j1\text{IV}}$ and $\hat{\delta}_{\text{PCIV}}$ are called PCIV estimators here, about which the reader may be confused. But the former estimates a linear combination of the coefficient vector δ, and the latter

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

estimates the coefficient vector δ itself by using $\hat{\theta}_{1\text{IIV}}$ and $\hat{\theta}_{2\text{IIV}}$. Because there does not seem to be a danger of confusion, we will continue to use the term *PCIV estimator* to refer to both $\hat{\theta}_{j\text{IIV}}$ and $\hat{\delta}_{\text{PCIV}}$.

12. Identifying relations (6) and (7) correspond with the facts that β cannot be estimated consistently under $\Pi_{11} = \Pi_{12} = 0$ and that γ can be.

13. Consider the case $\begin{bmatrix} \Pi_{11} \\ \Pi_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ \Upsilon \end{bmatrix}$, where Υ is of full column rank. Then, $\Gamma = \begin{bmatrix} S_{11}' \\ S_{12}' \end{bmatrix} = [0 \quad P]$ where P is a nonzero matrix, and $\hat{\theta}_{j\text{IIV}}$ do not contain any information on a certain part of β. Because Γ is of deficient column rank in this case, $\hat{\delta}_{\text{PCIV}}$ is not well defined.

14. See, for example, Bai and Perron (1998) for the least squares estimation of structural change points in linear regression.

15. Because the GDP and export growth rates are log differences of GDP and export, respectively, these can be regarded as stationary.

16. The optimal VAR lag for $(GDPR_i, EXPR_i)'$ is 2 according to the Bayesian information criterion.

17. The value of the sup-Wald test is 9.69.

18. Existing structural change tests cannot be used for the structural equation (28) because the equation may not be estimable under some data regimes.

REFERENCES

Andrews, D.W.K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821–856.

Anderson, T.W. & N. Kunitomo (1992) Tests of overidentification and predeterminedness in simultaneous equation models. *Journal of Econometrics* 54, 49–78.

Angrist, J.D. & A.B. Krueger (1995) Split sample instrumental variables estimates of the return to schooling. *Journal of Business and Economic Statistics* 13, 225–235.

Bai, J. & P. Perron (1998) Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47–78.

Bekker, P. (1994) Alternative approximations to the distributions of instrumental variables estimator. *Econometrica* 62, 657–681.

Bekker, P. & J. van der Ploeg (1999) Instrumental Variable Estimation Based on Grouped Data. Mimeo, Department of Economics, University of Groningen.

Bound, J., D.A. Jaeger, & R.M. Baker (1995) Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association* 90, 443–450.

Choi, I. & P.C.B. Phillips (1992) Asymptotic and finite sample distribution theory for IV estimators and tests in partially identified structural equations. *Journal of Econometrics* 51, 113–150.

Choi, I. & P.C.B. Phillips (1997) Regressions for Partially Identified, Cointegrated Simultaneous Equations. Cowles Foundation Discussion paper, Yale University.

Cragg, J.G. & S.G. Donald (1993) Testing identifiability and specification in instrumental variables models. *Econometric Theory* 9, 222–240.

Cragg, J.G. & S.G. Donald (1996) On the asymptotic properties of LDU-based tests of the rank of a matrix. *Journal of the American Statistical Association* 91, 1301–1309.

Dagenais, M.G. & D.L. Dagenais (1995) Higher moment estimators for linear regression models with errors in the variables. *Journal of Econometrics* 76, 193–221.

Dufour, J.-M. (1997) Some impossibility theorems in econometrics with applications to structural and dynamic models. *Econometrica* 65, 1365–1388.

Ghartey, E.E. (1993) Causal relationship between exports and economic growth: Some empirical evidence in Taiwan, Japan and the US. *Applied Economics* 25, 1145–1152.

Gill, L. & A. Lewbel (1992) Testing the rank and definiteness of estimated matrices with applications to factor, state-space, and ARMA models. *Journal of the American Statistical Association* 87, 766–776.

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

Hausman, J. (1983) Specification and estimation of simultaneous equation models. In Z. Griliches & M.D. Intriligator (eds.), *Handbook of Econometrics*, vol. 1. Amsterdam: North-Holland.
Hsiao, C. (1983) Identification. In Z. Griliches & M.D. Intriligator (eds.), *Handbook of Econometrics*, vol. 1. Amsterdam: North-Holland.
Judge, G.G. & M. Bock (1993) Biased estimation. In Z. Griliches & M.D. Intriligator (eds.), *Handbook of Econometrics*, vol. 1. Amsterdam: North-Holland.
Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lütekepohl, & T.-S. Lee (1985) *The Theory and Practice of Econometrics*. New York: Wiley.
Koopmans, T.C. & W.C. Hood (1953) The estimation of simultaneous linear economic relationships. In W.C. Hood & T.C. Koopmans (eds.), *Studies in Econometric Method*, Cowles Foundation Monograph 10. New York: Wiley.
Lee, F.-Y. & W.E. Cole (1994) Simultaneity in the study of exports and economic growth. *International Economic Journal* 8, 33–41.
Lewbel, A. (1997) Constructing instruments for regressions with mismeasurement error when no additional data are available with an application to patents and R&D. *Econometrica* 65, 1201–1213.
Lucas, G. (1976) Econometric policy evaluation: A critique. In K. Brunner & A.H. Meltzer (eds.), *Carnegie-Rochester Conference Series on Public Policy*, vol. 1. Amsterdam: North-Holland.
Mariano, R.S. (1982) Analytical small sample theory in econometrics: The simultaneous equations case. *International Economic Review* 23, 503–534.
Marin, D. (1992) Is the export-led growth hypothesis valid for industrialized countries? *Review of Economics and Statistics* 74, 678–688.
Muirhead, R. (1982) *Aspects of Multivariate Statistical Theory.* New York: Wiley.
Newey, W. (1990) Efficient instrumental variables estimation of nonlinear models. *Econometrica* 58, 809–837.
Newey, W. & D. McFadden (1994) Large sample estimation and hypothesis testing. In R. Engle & D.L. McFaddenr (eds.), *Handbook of Econometrics*, vol. 4. Amsterdam: North-Holland.
Perron, P. (1989) The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57, 1361–1401.
Phillips, P.C.B. (1983) Exact small sample theory in the simultaneous equations model. In Z. Griliches & M.D. Intriligator (eds.), *Handbook of Econometrics*, vol. 1. Amsterdam: North-Holland.
Phillips, P.C.B. (1989) Partially identified econometric models. *Econometric Theory* 5, 181–240.
Shan, J. & F. Sun (1999) Export-led growth and the US economy: Some further testing. *Applied Economics Letters* 6, 169–172.
Sims, C.A. (1980) Macroeconomics and reality. *Econometrica* 48, 1–48.
Staiger, D. & J.H. Stock (1997) Instrumental variables regression with weak instruments. *Econometrica* 65, 557–586.
Wang, J. & E. Zivot (1998) Inference on structural parameters in instrumental variables regression with weak instruments. *Econometrica* 66, 1389–1404.
Zivot, E., Startz, & Nelson (1998) Valid confidence intervals and inference in the presence of weak instruments. *International Economic Review* 39, 1119–1144.

APPENDIX: PROOFS

Proof of Theorem 1.

(a) Because $\hat{\pi} = \hat{\Pi}\delta + AU$, we may write

$$\hat{\delta}_{\mathrm{MD}} - \delta = (\hat{\Pi}'\Upsilon_{\delta}^{-1}\hat{\Pi})^{-1}\hat{\Pi}'\Upsilon_{\delta}^{-1}AU.$$

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

But Assumption 7 yields $\hat{\Pi} \xrightarrow{p} \begin{bmatrix} \Pi_{11} & I_k \\ \Pi_{12} & I_k \end{bmatrix}$,

$$NY_\delta = NAA' = \begin{bmatrix} \frac{N}{N_1}\left(\frac{1}{N_1} Z_{11}' Z_{11}\right)^{-1} & 0 \\ 0 & \frac{N}{N_2}\left(\frac{1}{N_2} Z_{12}' Z_{12}\right)^{-1} \end{bmatrix}$$

$$\xrightarrow{p} \begin{bmatrix} \frac{1}{d_1} M_{z_{11} z_{11}}^{-1} & 0 \\ 0 & \frac{1}{d_2} M_{z_{12} z_{12}}^{-1} \end{bmatrix}$$

and

$$\sqrt{N} AU \xrightarrow{d} N(0, \sigma_{uu} \text{plim}_{N\to\infty} NY_\delta),$$

from which the stated result follows.

(b) Using $\bar{\pi} = \bar{\Pi}\phi + BU$, write

$$\bar{\phi}_{\text{MD}} - \phi = (\bar{\Pi}' Y_\phi^{-1} \bar{\Pi})^{-1} \bar{\Pi}_\phi' Y_\phi^{-1} BU.$$

The result follows from the given assumptions, because $\bar{\Pi} \xrightarrow{p} \begin{bmatrix} \Pi_{11} & I_{k_1} & 0 \\ \Pi_{12} & I_{k_1} & 0 \\ \Pi_2 & 0 & I_{k_2} \end{bmatrix}$,

$$NY_\phi = NBB' \xrightarrow{p} \begin{bmatrix} \frac{1}{d_1} M_{z_{11} \bullet z_{21}}^{-1} & 0 & -M_{z_1 z_1}^{-1} M_{z_1 z_2} M_{z_2 \bullet z_1}^{-1} \\ 0 & \frac{1}{d_2} M_{z_{12} \bullet z_{22}}^{-1} & -M_{z_1 z_1}^{-1} M_{z_1 z_2} M_{z_2 \bullet z_1}^{-1} \\ -M_{z_2 \bullet z_1}^{-1} M_{z_2 z_1} M_{z_1 z_1}^{-1} & -M_{z_2 \bullet z_1}^{-1} M_{z_2 z_1} M_{z_1 z_1}^{-1} & M_{z_2 \bullet z_1}^{-1} \end{bmatrix}$$

and

$$\sqrt{N} BU \xrightarrow{d} N(0, \sigma_{uu} \text{plim}_{N\to\infty} NY_\phi).$$

Note that part (b) of Assumption 7 implies that $M_{z_{11} \bullet z_{21}}$ and $M_{z_{12} \bullet z_{22}}$ are nonsingular (cf. Muirhead, 1982, p. 586). Moreover, because $(1/N)Z'Z = Z^{1\prime}Z^1 + Z^{2\prime}Z^2$, the probability limit of $(1/N)Z'Z$ is a positive definite matrix by part (b) of Assumption 7. This implies that $M_{z_1 z_1}$ and $M_{z_2 \bullet z_1}$ are nonsingular.

Proof of Lemma 1.

(a) Let $H_j = [h_i]_{i \in \mathcal{G}_j}$. Because $(1/N_j)V_{2j}' H_j, (1/N_j)V_{2j}' Z_{1j} \xrightarrow{p} 0$ under Assumption 8, it follows that

$$\frac{1}{N_j} W_j' G_j \xrightarrow{p} \begin{bmatrix} \Pi_{1j}' M_{z_{1j} h_j} & \Pi_{1j}' M_{z_{1j} z_{1j}} \\ M_{z_{1j} h_j} & M_{z_{1j} z_{1j}} \end{bmatrix} = M_{W_j G_j},$$

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press

which shows that $\text{rank}(M_{W_j G_j}) = k_1$. Because $M_{G_j G_j}$ can be factorized as $M_{G_j G_j} = M_{G_j G_j}^{1/2} M_{G_j G_j}^{1/2}$ by Assumption 8, where $M_{G_j G_j}^{1/2}$ is of full rank, we have $\text{rank}(M_{W_j G_j} M_{G_j G_j}^{-1/2}) = k_1$, which completes the proof.

(b) Using the same methods as for part (a) yields the required result. ■

LEMMA 3. *Under Assumption 8,*

(a) $(1/N_j)\lambda_{ji} \xrightarrow{p} \xi_{ji}$ $(i = 1, \ldots, n + k_1)$, *where* ξ_{ji} *is as defined in Lemma 2.*

(b) $S_{j1} \xrightarrow{p} \Psi_{j1}$, *where* Ψ_{j1} *is a matrix whose columns are eigenvectors of the matrix* $M_{W_j G_j} M_{G_j G_j}^{-1} M_{G_j W_j}$ *corresponding to* k_1 *largest eigenvalues.*

Proof.

(a) λ_{ji} satisfies the determinantal equation $|(1/N_j) W_j' P_{G_j} W_j - (1/N_j)\lambda_{ji} I| = 0$. But $(1/N_j) W_j' P_{G_j} W_j \xrightarrow{p} M_{W_j G_j} M_{G_j G_j}^{-1} M_{G_j W_j}$, from which the result follows.

(b) The matrix S_{j1} satisfies the equation $(1/N_j) W_j' P_{G_j} W_j S_{j1} = (1/N_j)\lambda_{ji} S_{j1}$ and $(1/N_j) W_j' P_{G_j} W_j \xrightarrow{p} M_{W_j G_j} M_{G_j G_j}^{-1} M_{G_j W_j}$. Thus, we obtain the stated result by using part (a) of Lemma 3. ■

Proof of Lemma 2.

(a) Because $S_{j1}' S_{j2} = 0$, it follows that $\hat{\theta}_{j1\text{IV}} = S_{j1}'\delta + (X_{1j}' P_{G_j} X_{1j})^{-1} X_{1j}' P_{G_j} U_j$. But Lemma 3 gives

$$\frac{1}{N_j} X_{1j}' P_{G_j} X_{1j} = \frac{1}{N_j} S_{j1}' W_j' G_j (G_j' G_j)^{-1} G_j' W_j S_{j1}$$

$$= \frac{1}{N_j} \text{diag}(\lambda_{j1}, \ldots, \lambda_{jk_1}) \xrightarrow{p} \text{diag}(\xi_1, \ldots, \xi_{k_1}), \tag{A.1}$$

and, as a result of Assumption 8,

$$\frac{1}{\sqrt{N_j}} X_{1j}' P_{G_j} U_j = \frac{1}{\sqrt{N_j}} S_{j1}' W_j' G_j (G_j' G_j)^{-1} G_j' U_j \xrightarrow{d} N(0, \sigma_{uu} \text{diag}(\xi_1, \ldots, \xi_{k_1})). \tag{A.2}$$

Relations (A.1) and (A.2) yield the stated result.

(b) This is obtained by using the same methods as for part (a). ■

Proof of Theorem 2. Choose a matrix C_j such that $C_j' C_j = N_j M_{W_j G_j} M_{G_j G_j}^{-1} M_{G_j W_j}$. At $N_j = \infty$, $\sqrt{N_j}(\hat{\theta}_{j1\text{IV}} - S_{11}'\delta)$ has the same distribution as $(C_j' C_j)^{-1} C_j' U_j$. But $E[(C_1' C_1)^{-1} C_1' U_1 U_2' C_2' (C_2' C_2)^{-1}] = 0$ for any choice of the matrices C_1 and C_2. Therefore, $\sqrt{N_1}(\hat{\theta}_{11\text{IV}} - S_{11}'\delta)$ and $\sqrt{N_2}(\hat{\theta}_{21\text{IV}} - S_{21}'\delta)$ are independent at $N_1, N_2 = \infty$. Now, writing

$$\sqrt{N}(\hat{\delta}_{\text{PCIV}} - \delta) = (\Gamma' \Theta_\delta^{-1} \Gamma)^{-1} \Gamma' \Theta_\delta^{-1} \begin{bmatrix} \sqrt{N/N_1}\sqrt{N_1}(\hat{\theta}_{11\text{IV}} - S_{11}'\delta) \\ \sqrt{N/N_2}\sqrt{N_2}(\hat{\theta}_{21\text{IV}} - S_{21}'\delta) \end{bmatrix}$$

and using part (a) of Lemma 2, part (b) of Lemma 3, and (A.1) completes the proof for part (a). Part (b) can be shown in a similar manner. ■

https://doi.org/10.1017/S0266466602183095 Published online by Cambridge University Press